ELSEVIER

Biological Psychology 46 (1997) 143–168

BIOLOGICAL PSYCHOLOGY

Brain potentials elicited by words: word length and frequency predict the latency of an early negativity

Lee Osterhout [a,*], Michael Bersick [a], Richard McKinnon [b]

[a] *Department of Psychology, University of Washington, Seattle, WA 98195, USA*
[b] *Department of Linguistics, University of Washington, Seattle, WA 98195, USA*

Received 7 November 1995; received in revised form 14 January 1997; accepted 14 January 1997

Abstract

Prior work has suggested that open- and closed-class words elicit negative components in the event-related potential (ERP) that differ in timing and scalp distribution. We tested this hypothesis against the possibility that the word-class effects are attributable to quantitative differences in word length and frequency. Event-related brain potentials (ERPs) were recorded from 13 scalp sites while participants read normal or scrambled prose. ERPs were averaged as a function of word class (open vs. closed) and grammatical category (articles, nouns, verbs, etc.). Regression analyses indicated that the latency of an early-occurring negative component was highly correlated with the mean length and normative frequency of words in each grammatical category. Stronger correlations were observed in the scrambled prose condition than in the normal prose condition. Differences in the scalp distributions of these negativities were found to be a function of grammatical category rather than word class. These results are taken to be inconsistent with the claim that open- and closed-class words elicit qualitatively distinct negativities. © 1997 Elsevier Science Ireland Ltd.

Keywords: Event-related potentials; Word class; Language comprehension

---

* Corresponding author. Tel.: + 1 206 5434177; e-mail: losterho@u.washington.edu.

PII S0301-0511(97)05250-2

1. Introduction

In English, the distinction between syntax (sentence form) and semantics (sentence meaning) is roughly paralleled by the distinction between two lexical categories, the closed-class and open-class vocabularies (Garrett, 1982). Closed-class words (e.g. articles, auxiliary verbs, prepositions, and pronouns) typically act as agents of phrasal construction, whereas open-class words (e.g. nouns, verbs, and adjectives) convey meaning by referring to specific objects and events. Studies of aphasic syndromes have suggested that the use of these two word classes during language processing can be selectively disrupted (c.f. Garrett, 1982; Zurif and Swinney, 1994). For example, lesions to anterior regions of the left hemisphere can disrupt the production and comprehension of closed-class words, whereas lesions to posterior regions disrupt use of open-class words (Caramazza and Zurif, 1976; Damasio and Damasio, 1989; Heilman and Scholes, 1976; Saffran et al., 1980; Schwartz et al., 1980). Research with normal adults has indicated that the two word classes participate in distinct types of speech errors (Garrett, 1980) and display different patterns of hemispheric asymmetries (Bradley and Garrett, 1983). Such evidence has been taken to indicate that open- and closed-class words might be both functionally and neuroanatomically distinct.

A particularly compelling type of evidence would be neurobiological evidence from the normal brain that the two word classes are distinct. One means for obtaining such evidence involves the recording of event-related brain potentials (ERPs) elicited during language processing. ERPs are a continuous, on-line record of the brain's electrical activity that provides at least rough estimates of lateralization and localization of function (c.f. Coles and Rugg, 1995). ERPs consist of positive and negative voltage peaks (or 'components') that are distributed over time. It is generally agreed that ERP components with distinct scalp distributions are necessarily generated by distinct neural systems (c.f. Johnson, 1993). Therefore, evidence that open- and closed-class words elicit different ERP components, or different distributions of components across the scalp, would support the notion that the two word classes are neurally (and, by extension, cognitively) distinct.

Evidence of exactly this type has been reported by Neville et al. (1992). These researchers recorded ERPs while participants read a set of unrelated sentences. Neville et al. noted several differences in the ERPs to open- and closed-class words. Some of these differences had been noted in other reports (e.g. Kutas and Hillyard, 1983). For example, closed-class words elicited a ramp-like negativity between about 400 and 700 ms after presentation of the word of interest (labeled the *N400-700* effect), whereas open-class words elicited a late positive component (LPC) in this window (for a review, see Kutas and Van Petten, 1994)

Of particular note, however, was Neville et al.'s observation that the two word classes elicited negative components that clearly differed in their temporal and spatial properties. Closed-class words elicited a negative component that was largest over anterior regions of the left hemisphere and that peaked at about 280 ms (*N280*). By contrast, open-class words elicited a large-amplitude, posteriorly distributed negativity that peaked between 350 and 400 ms (*N400*)[1]. This finding is

---

[1] Closed-class words also elicited an N400-like component, but this effect was much smaller in amplitude and more restricted in distribution than the N400 elicited by open-class words.

interesting for several reasons. Firstly, the distribution of these negativities corresponds quite well with the clinical observation that left anterior lesions disrupt use of closed-class words whereas posterior lesions disrupt use of open-class words. Secondly, Neville et al. repeated their experiment using deaf participants who had acquired American Sign Language (ASL) but who had not acquired full competence with English grammar. Although open-class words elicited a similar N400 in both hearing and deaf populations, the closed-class words elicited an N280 only in the hearing (i.e. grammatically competent) participants. Finally, a large body of literature has demonstrated that the N400 component is primarily sensitive to semantic aspects of language processing (Kutas and Hillyard, 1980a,b,c; for a review, see Kutas and Van Petten, 1994 or Osterhout and Holcomb, 1995), whereas other studies have shown that certain types of syntactic violations elicit a left anterior negativity between roughly 200 and 400 ms (the *LAN* effect; c.f. Kutas and Van Petten, 1994). Although the relationship between the LAN effect and the N280 remains obscure, these results are consistent with the notion that syntactic processing involves left anterior regions of cortex whereas semantic processing involves posterior regions.

These results seem to provide compelling support for the hypothesis that qualitative differences exist in the representation and function of open- and closed-class words and that ERPs are sensitive to these differences. Nonetheless, there are at least two issues that need to be resolved before this conclusion can be accepted with confidence. Firstly, not all researchers have observed an N280-like negativity in the ERP response to closed-class words (Kutas and Hillyard, 1983; Kutas et al., 1988; Pulvermüller et al., 1995), and in several studies the two word classes have elicited negative-going components that were quite similar in timing and/or distribution (Garnsey, 1985; Nobre and McCarthy, 1994; Van Petten and Kutas, 1991). For example, Garnsey (1985) reported that ERPs to the two word classes were highly similar and found no effects or interactions involving word class. Kutas and Hillyard (1983) found that although closed-class words elicited a negative-going peak at about 300 ms, this effect was localized to anterior but not left-anterior sites. Pulvermüller et al. (1995) reported that closed-class words elicited negativities that peaked at about 160 and 200 ms, and that these effects were larger over the left hemisphere. However, open-class words also elicited negativities at similar latencies, although these effects were more symmetrical in distribution. These disparate results might partly be due to wide variation across experiments in the type of stimuli presented and in the task assigned to participants. Stimuli have included prose (Kutas and Hillyard, 1983), a series of unrelated sentences (Neville et al., 1992; Van Petten and Kutas, 1991), and lists of unrelated words (Garnsey, 1985; Nobre and McCarthy, 1994; Pulvermüller et al., 1995). Tasks have included reading for comprehension (Kutas and Hillyard, 1983), sentence-acceptability judgments (Neville et al., 1992), and lexical decisions (Pulvermüller et al., 1995).

A second reason for caution in interpreting the Neville et al. results revolves around the difficulty of determining which aspects of the vocabulary distinction (word length, frequency of usage, repetition, contextual constraints, abstractness of meaning, referentiality, grammatical role, etc.) are responsible for the reported ERP differences (c.f. Kutas and Van Petten, 1994). In particular, it is critical that the effects of *quantitative* differences between words (e.g. word frequency and length) are not mistaken for the effects of *qualitative* differences between word classes. Although some studies have carefully controlled for the factors of word frequency and length (Garnsey, 1985; Pulvermüller et al., 1995), others have not (e.g. Kutas and Hillyard, 1983; Van Petten and Kutas, 1991). Neville et al. examined the effects of word frequency and length by forming separate averages for high and low frequency and short and long open-class words. None of these categories elicited an N280-like effect, but all categories elicited an N400. Furthermore, the primary effect of word frequency was to alter N400 amplitude (such that less-frequent words elicited larger N400s; see also Van Petten and Kutas, 1991), whereas the major effect of word length was to add a symmetrical frontal positivity in the ERPs to short words, beginning at about 200 ms and continuing for several hundred ms. Based on these results, Neville et al. concluded that the N280/N400 distinction was in fact a function of word class rather than word frequency or length. However, as noted by Pulvermüller et al. (1995), this procedure might not reveal important frequency effects due to the fact that the highest frequency range is occupied almost exclusively by function words (Francis and Kucera, 1982). Correspondingly, variations in frequency would tend to be larger across word class than among members of the open class. Certain frequency-related effects might be observable only under conditions in which the frequency range is sufficiently wide.

The research reported here had two goals. Firstly, we observed the ERP response to open- and closed-class words under two stimulus conditions (normal prose and scrambled prose) in order to directly assess whether or not ERP differences related to word class are robust across different stimulus conditions. This manipulation also allowed us to determine, to a first approximation, whether these ERP effects reflect *word-level* processing or *sentence-level* processing (c.f. Neville et al., 1992). The N280/N400 distinction might reflect events occurring at the lexical level, e.g. some aspect of word recognition or lexical access[2]. Alternatively, these effects might reflect events occurring at the *sentence* or *discourse* level. For example, the N280 might reflect processes associated with the structural analysis of sentences, whereas the N400 might be associated with the semantic integration of words into the ongoing discourse (c.f. Neville et al., 1992; Osterhout and Holcomb, 1995). Because text randomization does not affect the lexical properties of words, ERP effects that primarily reflect word processing should be observed in both prose conditions. By contrast, scrambled prose is not constructed of sentences and should therefore not engage the processes involved in sentence processing; hence, ERP effects that primarily reflect sentence processing should not be observed when participants are asked to read scrambled prose.

---

[2] Recent work has been cited as evidence that the amplitude of the N400 component is not uniquely or even primarily sensitive to processes underlying word recognition and lexical access (e.g. Brown and Hagoort, 1993; Holcomb, 1993; for discussion, see Osterhout and Holcomb, 1995). Also, the observation that closed-class words elicit an N400 component, albeit much smaller in amplitude than that elicited by open-class words, might be viewed as problematic for the notion that the N280/N400 distinction reflects the accessing of separable lexicons for the two word classes.

The second goal was to carefully examine the effects of word frequency and length on ERPs to words. Rather than controlling these factors by equating them across word class as has been done by some researchers (c.f. Garnsey, 1985), we presented prose with normal variation in word frequency and length. By forming averages over grammatical categories (article, preposition, noun, verb, etc.) and determining the mean lengths and frequencies of words in each category, we were able to perform regression analyses correlating these factors with relevant changes in the ERPs. This procedure also allowed us to observe possible differences in the ERPs elicited by the various grammatical categories. Theories of lexical semantics propose that the grammatical categories are characterized by lexical conceptual structures of varying semantic richness (e.g. Jackendoff, 1983; 1990; for a discussion, see Flores d'Arcais, 1985), and recent research with normal (Friederici, 1985) and aphasic populations (Jarema and Friederici, 1994) points to the importance of distinguishing between different categories of function words. Hence, ERPs averaged as a function of word class might obscure the effects of grammatical category.

2. Experiment 1

2.1. Method

2.1.1. Participants

Fourteen right-handed, native-English-speaking undergraduates (seven females and seven males) participated for course credit. Ages ranged from 18 to 20 (mean = 19) years.

2.1.2. Materials

A short essay describing Amelia Earhart's last flight was selected for presentation. This text was a modified version of an essay taken from an English as a Second Language text (Kenan, 1986). The text contained 173 sentences that ranged from four to 15 words in length. ERPs to sentence-initial and sentence-final words were not included in the grand average waveforms. This was done to avoid confounding the effects of word class with presumed processes that are specific to the sentence-initial and sentence-final positions (e.g. sentence wrap-up processes). In most cases, the sentences began with a closed-class word and ended with an open-class word. The remaining text contained 746 closed-class words (including 120 articles, 130 prepositions, 93 pronouns, and 146 auxiliary verbs) and 484 open-class words (including 199 nouns and 180 verbs).

2.1.3. Procedure

Participants sat comfortably in an isolated, dimly illuminated room. The text was presented as a series of sentences. Each sentence trial consisted of the following events: a fixation cross appeared for 500 ms, after which a sentence was presented in a word-by-word manner, with each word appearing on the center of the screen for 300 ms. A blank-screen interval of 400 ms separated words. Sentence-ending words appeared with a period. A 1450-ms blank-screen interval followed each sentence and provided participants with an opportunity to blink and rest their eyes. This interval was followed by a prompt asking participants to respond by pressing a button on a joystick when ready for the next sentence. The hand used to respond was counterbalanced across participants. Participants were instructed to read the text carefully for comprehension and to be prepared to answer questions about the content of the text at the end of the experiment. Each session lasted approximately 2 h.

2.1.4. Data acquisition and analysis

Continuous EEG was recorded from 13 scalp sites using tin electrodes attached to an elastic cap (Electrocap International). Electrode placement included International 10–20 system locations (Jasper, 1958) over homologous positions over the left and right occipital (O1, O2) and frontal (F7, F8) regions and from frontal (Fz), central (Cz), and parietal (Pz) midline locations. In addition, several nonstandard sites over posited language centers were used, including Wernicke's area and its right hemisphere homologue (WL, WR: 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz), posterior temporal (TL, TR: 33% of the interaural distance lateral to Cz), and anterior temporal (ATL, ATR: one-half the distance between F7/F8 and T3/T4). Vertical eye movements and blinks were monitored by means of two electrodes, one placed beneath the left eye and another placed to the right of the right eye. The above 15 channels were referenced to an electrode placed over the left mastoid bone and were amplified with a bandpass of 0.01 to 100 Hz (3 db cut-off) by a Grass Model 12 amplifier system. Activity over the right mastoid was actively recorded on a sixteenth channel to determine if there were any effects of the experimental variables on the mastoid recordings. No such effects were observed.

Continuous analog-to-digital conversion of the EEG and stimulus trigger codes was performed by a Data Translation 2801-A board and a 486-based computer at a sampling frequency of 200 Hz. Epochs were comprised of the 100 ms preceding and the 1180 ms following presentation of individual words in the sentences. Trials characterized by excessive eye movement or amplifier blocking were removed prior to averaging (approximately 10 and 20% of the trials were removed due to artifact in Experiments 1 and 2, respectively).

ERP components of interest were quantified by computer as either mean voltage or peak (maximal) amplitude within a latency range, relative to the 100 ms of activity immediately preceding presentation of the word of interest. For measurements of mean voltage, windows between 50–150, 150–250, 250–350, 350–450, and 450–700 ms were used. These windows roughly correspond to the latency ranges of the N1, P2, N280, N400, and N400–700/LPC components described in previously published reports. Additional analyses are described when reported. Because the primary questions of interest are most directly addressed by data acquired over lateral sites, and in an effort to reduce data analysis, analyses are reported only on data from lateral electrode sites[3]. ERP measures from all lateral

---

[3] Analyses were also performed on data from midline sites but were not reported. These analyses were fully consistent with those reported for the lateral sites.

sites were subjected to a three-way ANOVA with repeated measures on word type, two levels of hemisphere, and five levels of electrode site. The Greenhouse and Geisser (1959) correction was applied to all repeated measures with greater than one degree of freedom in the numerator. In such cases, the corrected P-value is reported. For analyses involving significant interactions between sentence types and electrode sites in the presence of a reliable main effect of sentence type, two sets of analyses were reported: analyses on raw data, and analyses on data that have been normalized following the procedure described by McCarthy and Wood (1985). This normalization procedure was used because in certain cases spurious interactions can result if the experimental effects are of different overall amplitude. In order to minimize the number of reported analyses, analyses on normalized data are reported only when the effect is deemed to be theoretically important (i.e. when the distributions of the responses to two conditions are directly compared).

3. Results

3.1. Averages over word class

Fig. 1A plots ERPs (recorded over the 10 lateral scalp sites) elicited by open- and closed-class words. Fig. 1B overplots ERPs from the left and right hemispheres for each word class. Inspection of Fig. 1 reveals that words elicited a series of positive and negative voltage deflections similar to those reported previously. A clear negative-positive complex was visible in the first 300 ms after stimulus presentation. The negative component ('N1') peaked at about 120 ms. The positive component ('P2') peaked at about 200 ms and, for both word class conditions, was larger in amplitude over anterior than over posterior regions (electrode site: $F(4,52) = 11.25$, $P < 0.003$, MSe $= 3.70$) and larger over the right than the left hemisphere, particularly at posterior sites (hemisphere × electrode site: $F(4,52) = 4.09$, $P < 0.03$; MSe $= 0.50$).

As in prior reports (Kutas and Hillyard, 1983; Kutas and Van Petten, 1994; Neville et al., 1992), ERPs to open-class words were more negative-going than those to closed-class words beginning at about 200 ms. No reliable differences were found in the 50–150 ms window. However, ERPs to closed-class words were more negative-going than those to open-class words within both the 150–250 ms ($F(1,13) = 29.28$, $P < 0.001$, MSe $= 1.99$) and 250–350 ms windows ($F(1,13) = 15.15$, $P < 0.01$, MSe $= 1.98$). No reliable differences between word classes were observed within the N400 window (350–450 ms), $P > 0.1$. Between 450 and 700 ms, ERPs to closed-class words were negative in voltage (N400-700) whereas ERPs to open-class words were positive in voltage (LPC) ($F(1,13) = 88.23$, $P < 0.0001$, MSe $= 2.14$. These differences were largest over temporal sites in both hemispheres (word class × electrode site: raw data, $F(4,52) = 7.98$, $P < 0.01$, MSe $= 0.23$; normalized data: $F(4,52) = 8.13$, $P < 0.01$). In general, these word class differences are quite similar to those reported previously by other investigators (e.g. Kutas and Hillyard, 1983).

However, even though ERPs to closed-class words were reliably more negative-going than were those to open-class words within the 'N280' window, i.e. between 250 and 350 ms, closed-class words did not elicit a clearly defined N280 component similar to that reported by Neville et al. (1992). Instead, closed-class words elicited two negative-going effects (in addition to the more broadly distributed N400–700) that were not elicited by open-class words. One effect, most visible at temporal sites, peaked at about 350 ms before becoming more positive-going at about 400 ms. Although this effect did not have a clearly defined morphology or peak, we will refer to it as the *N350*. A second, more clearly defined component peaked at about 450 ms, most notably over posterior sites (N450)[4]. Peak latencies and amplitudes were then computed in order to compare the latencies and scalp topographies of the N350 and N450 elicited by closed-class words to the latency and topography of the N400 component elicited by open-class words. Peak latencies and amplitudes were determined by finding the most negative-going point within latency windows of 300–400, 400–500, and 300–450 ms for the N350, N450, and N400 effects,

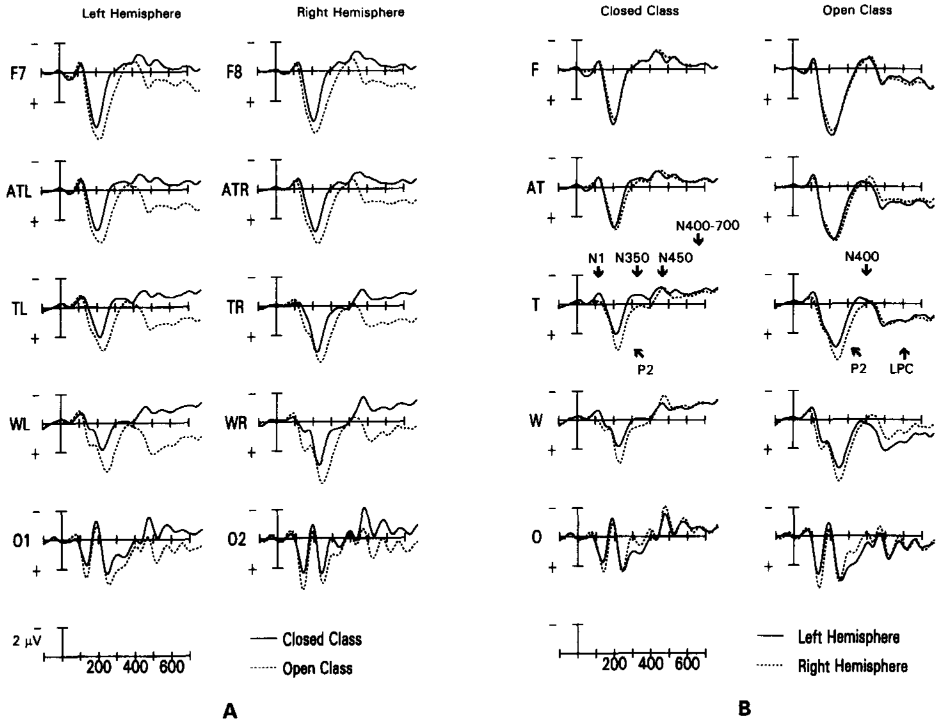

Fig. 1. Grand average ERPs recorded over 10 lateral sites to closed-class and open-class words in the normal prose condition. Each hashmark represents 100 ms. Positive voltage is plotted down. 1A overplots activity elicited by open- and closed-class words. 1B overplots activity recorded over the left and right hemispheres.

---

[4] Given the rate of stimulus presentation (each word displayed for 300 ms), the N450 might be caused (in part or entirely) by stimulus offset effects.

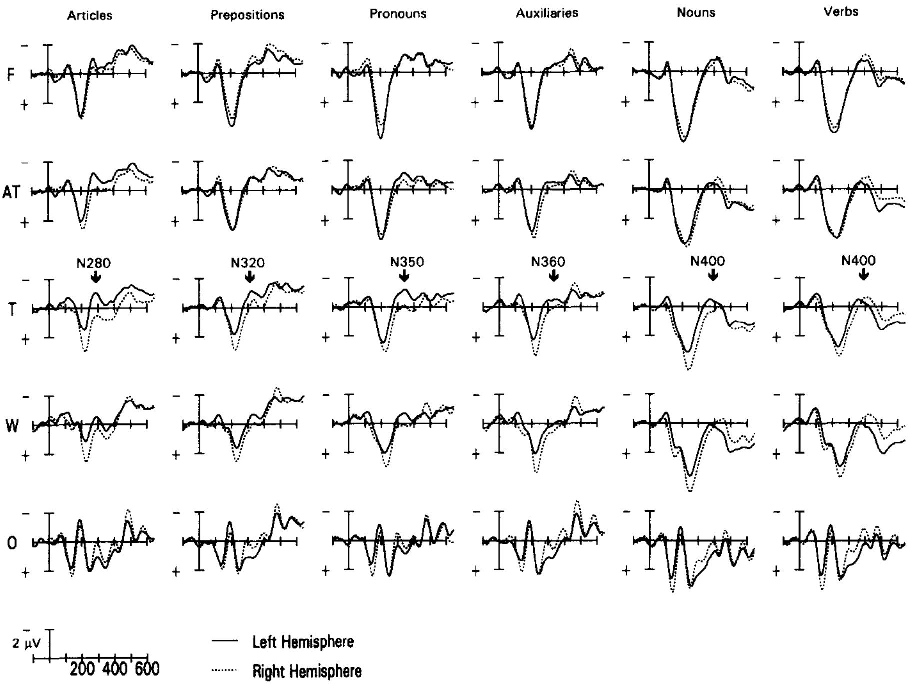

Fig. 2. ERPs averaged as a function of grammatical category for words in the normal prose condition.

respectively. N400 latency was reliably later-occurring than N350 latency, $F(1,13) = 11.30$, $P < 0.01$, $MSe = 0.003$, but earlier-occurring than N450 latency, $F(1,13) = 40.06$, $P < 0.0001$, $MSe = 0.009$. Evidence of distinct scalp distributions for these negativities would take the form of a three-way interaction between word class, hemisphere, and electrode site. Analyses on peak amplitudes revealed highly unreliable interactions (N350 vs. N400: $F(4,52) = 1.05$, $P > 0.4$; N450 vs. N400: $F(4,52) = 1.23$, $P > 0.3$).

3.2. Averages over grammatical categories

Words were then subdivided on the basis of grammatical category. Grand-average waveforms were computed for all grammatical categories that contained 90 or more exemplars within the text shown to participants. This criterion was chosen as it provided an adequate signal-to-noise ratio for examining the expected effects of interest. These subclasses included four types of closed-class words (articles, prepositions, pronouns, and auxiliary verbs) and two types of open-class words (nouns and verbs)[5].

---

[5] We also examined ERPs elicited by the open-class category of adjective. This category contained 74 exemplars. ERPs to adjectives were more variable than were those to the other categories (due in part to the smaller number of exemplars in the adjective category) but were quite similar in morphology, latency, and distribution to those elicited by the two other open-class categories (nouns and verbs).

ERPs averaged over exemplars in each category are shown in Fig. 2. All categories elicited a negative-going component that peaked between 280 and 400 ms, although this component was less robust for auxiliaries than for the other categories. The peak latency of the negativity varied across word type. The peak was earliest-occurring for articles (N280), intermediate for prepositions (N320), pronouns (N350), and auxiliary verbs (N360), and latest-occurring for nouns and verbs (N400). For all closed-class grammatical categories, hemispheric asymmetries (left greater than right) were largest over temporal and Wernicke's sites. For articles, these asymmetries extended to the anterior temporal sites. For nouns and verbs, ERPs tended to be more positive-going in the right hemisphere until about 400 ms, at which point ERPs were more positive-going in the left hemisphere. This shift reflected the fact that the N400 component peaked slightly later and was slightly larger in the right than in the left hemisphere.

Statistical analyses were performed to determine whether these negativities differed in scalp distribution and latency. To this end, peak amplitudes and latencies were determined within latency windows of 250–360 ms for closed-class words and 250–450 ms for open-class words, over all lateral electrode sites[6]. Analyses on peak amplitude revealed marginally reliable differences in scalp distribution across categories (grammatical category × hemisphere: raw data, $F(5,65) = 2.42$, $P = 0.07$, MSe = 1.29; normalized data, $F(5,65) = 2.63$, $P = 0.06$). Selected planned comparisons were then carried out on peak amplitudes involving three comparisons within the closed class categories (involving articles, prepositions, and pronouns), one comparison involving categories within the open class, and three between-class comparisons (comparing articles, prepositions, and pronouns to nouns)[7]. Differences in scalp distribution were tested using a modified Bonferroni correction for multiple comparisons (Keppel, 1982). Under this procedure, we used an alpha level of 0.035 for each comparison. Comparisons within the closed-class categories revealed reliable distributional differences only between articles and prepositions (category × hemisphere: $F(1,13) = 9.82$, $P < 0.01$, MSe = 0.70; category × hemisphere × electrode site: $F(4,52) = 4.45$, $P < 0.01$, MSe = 0.17). The two categories from the open class (verbs and nouns) did not reliably differ in scalp distribution. Between-class comparisons revealed reliable distributional differences between articles and nouns (category × hemisphere: $F(1,13) = 5,50$, $P = 0.035$, MSe = 1.43) and between prepositions and nouns (category × electrode site: $F(4,52) = 3.60$, $P < 0.01$, MSe = 0.41).

---

[6] We should note that Neville et al. (1992) reported that the N280 was present only (or primarily) over anterior regions of the left hemisphere. However, the negativity we observed was notable at most recording sites within both hemispheres, although of varying amplitude across sites and hemispheres. We therefore included data from all lateral sites in our analyses involving peak latency and amplitude rather than restricting our analyses to specific regions of the scalp.

[7] We performed selected comparisons in order to reduce the number of reported analyses and to minimize the necessary correction for familywise error.

Differences in latency across conditions were highly reliable $F(5,65) = 33.30$, $P < 0.001$, MSe = 0.004. In order to determine the relationship between latency and the lexical properties of word frequency and length, we computed mean normative frequencies and word lengths in letters for words in each grammatical category. Each participant's mean negative peak latency over all lateral electrode sites was computed for words in the six grammatical categories and was regressed onto models using as single predictors mean length and the logarithm of the mean normative frequency. Fig. 3 shows the resulting regression equations and scatterplots using the predictors of length (top panel) and frequency (bottom panel). The robustly significant correlations (length: $F(1,82) = 70.64$, $P < 0.0001$; frequency: $F(1,82) = 71.58$, $P < 0.0001$) indicated that length and frequency each accounted for

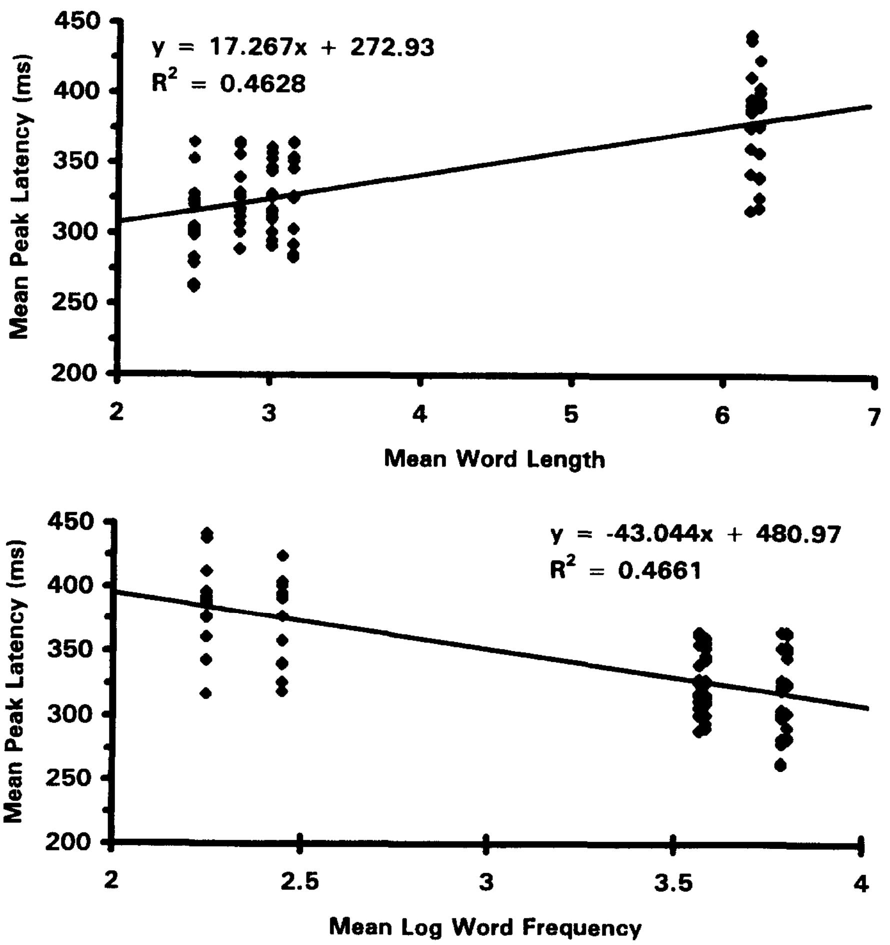

Fig. 3. Regression analyses correlating peak latency of the early-occurring negativity (in ms) with the mean length (top panel) and the log of mean normative frequency (bottom panel) of words in the normal prose condition.

Table 1
Individual regression parameters (ms) and proportion of variance accounted for, for mean peak latency regressed onto mean word length and mean log word frequency, normal text

Subject	Length			Log frequency		
	Intercept (ms)	Slope (ms)	R^2	Intercept (ms)	Slope (ms)	R^2
1	327	11	0.97	466	−28	0.91
2	259	18	0.73	490	−48	0.82
3	210	30	0.93	559	−71	0.82
4	289	5	0.36	348	−12	0.38
5	304	14	0.82	480	−36	0.84
6	268	26	0.96	584	−65	0.96
7	269	20	0.99	502	−48	0.98
8	245	21	0.69	508	−56	0.79
9	289	10	0.81	405	−23	0.70
10	261	14	0.79	442	−38	0.92
11	238	32	0.76	620	−79	0.74
12	258	12	0.76	413	−33	0.86
13	264	18	0.77	480	−45	0.76
14	339	8	0.91	436	−20	0.86

about 46% of the variance. Because length and log frequency were highly negatively correlated (-0.98), a model including both factors as predictors accounted for no additional variance. The intercept, slope, and R^2-values for individual participants are shown in Table 1. Inspection of Table 1 reveals that the R^2-values for individual subjects were, with one exception, larger in magnitude than the R^2-value (0.46) obtained by averaging over all subjects. This reflected the fact that the between-subject variance (MSe = 0.0034) was greater than the within-subject variance (MSe = 0.0013), $P < 0.01$.

Finally, inspection of Fig. 2 reveals slight differences in the anterior-posterior scalp distribution of the N400–700 across grammatical categories within the closed class. To determine the reliability of these differences, an ANOVA was performed on mean amplitude between 450 and 700 ms elicited by each of the closed class categories. This analysis revealed a reliable interaction between category and electrode site, but only when the data were not normalized (raw data: $F(12,156) = 2.77$, $P < 0.05$, MSe = 0.53; normalized data: $F(12,156) = 1.81$, $P < 0.1$).

4. Discussion

The results of Experiment 1 do not support the claim that the N280 component is a categorical marker of the closed class and, more generally, that early-occurring negativities reflect qualitative differences in the processing of the two word classes. As in previous reports (cf Kutas and Hillyard, 1983; Kutas and Van Petten, 1994; Neville et al., 1992), averages formed as a function of word class revealed that open- and closed-class words elicited reliably different ERPs. Open-class words elicited a larger-amplitude P2, a negativity peaking at about 400 ms (N400), and a subsequent late positive component (LPC). Closed-class words elicited two negative-going effects (N350 and N450) as well as a more sustained negativity (N400–700). Critically, however, closed-class words as a category failed to elicit a left-lateralized anterior negativity with a peak at about 280 ms (N280), and the negativities that were elicited by closed-class words (N350 and N450) had scalp distributions similar to that of the N400 elicited by open-class words.

Furthermore, averages formed over grammatical categories revealed that each category elicited a negative-going effect that immediately followed the N1-P2 complex. The latencies of these negativities were highly correlated with the lexical properties of word length and frequency. Latency variation in these negativities therefore reflected quantitative differences in lexical properties rather than qualitative differences corresponding to the open- and closed-class distinction. There were also some reliable differences in the scalp distributions of these negativities. However, these differences appeared to be a function of grammatical category rather than word class and were restricted to categories that were farthest apart in terms of the latency of the negativity.

One caveat that should be noted here is the ambiguity introduced by the possible effects of component overlap. For example, reliable differences between word classes and hemispheric asymmetries were observed within the 'N280 window,' i.e., between 250 and 350 ms. These effects were similar in some respects to those reported by Neville et al. (1992). Furthermore, several reliable differences in scalp distribution were observed when comparing peak amplitudes across grammatical categories. However, due to the complications introduced by component overlap, it is impossible to know for certain whether these effects reflect the hemispheric differences caused by the right-lateralized P2 component or, conversely, shifts in the distribution of the subsequent negativities across conditions. Although we do not have a resolution for this problem, we would note that this ambiguity of interpretation also exists with respect to the data reported by Neville et al. (1992).

A second caveat is that some or all of these conclusions might pertain only to the experimental conditions of Experiment 1, namely, conditions under which participants read lengthy, coherent prose. In most previous reports, participants have read a series of unrelated sentences (e.g. Neville et al., 1992; Van Petten and Kutas, 1991) or lists of unrelated words (e.g. Garnsey, 1985; Nobre and McCarthy, 1994; Pulvermüller et al., 1995; Van Petten and Kutas, 1991). The type of stimulus presented has been observed to have differential effects on ERPs to open- and closed-class words. For example, Van Petten and Kutas (1991) asked participants to read semantically meaningful, syntactically legal but nonsensical, and random word strings. N400 amplitude to open-class words was reduced in the semantically meaningful condition relative to the other two conditions, whereas N400 amplitude to closed-class words was reduced in both the semantically meaningful and syntactically legal but nonsensical conditions. In Experiment 2, we examined the influence of stimulus type by randomizing the prose passage presented in Experiment 1. The resulting scrambled prose was similar to the 'random word string' condition employed by Van Petten and Kutas.

Finally, as noted above, the normal prose/scrambled prose contrast allowed us to investigate whether these ERP effects reflect *word-level* processing or *sentence-level* processing. ERP effects observed in Experiment 1 that reflect word processing should be observed even when scrambled prose is presented. By contrast, ERP effects that reflect sentence processing should not be observed when participants read scrambled prose that contains no sentence structure.

5. Experiment 2

5.1. Method

5.1.1. Participants

Fourteen right-handed native-English-speaking undergraduates (eight females, six males) participated for course credit. None had participated in Experiment 1. Ages ranged from 18 to 29 (mean = 21) years.

5.1.2. Materials

The text presented in Experiment 1 was split into two halves. The words in each half were then randomized and divided into sentence-sized chunks such that the length of each 'chunk' was matched to the length of a sentence presented in Experiment 1. We presented 'pseudosentences' rather than uninterrupted lists of words for two reasons. Firstly, we wanted to make the stimuli as similar as possible to those presented in Experiment 1, the primary difference being the presence (Experiment 1) or absence (Experiment 2) of syntactic structure. Secondly, the blank-screen interval between the pseudosentences provided participants with an opportunity to blink and rest their eyes. We removed all punctuation and uppercase letters (excepting those in proper names).

5.1.3. Procedures

Procedures were identical to those employed during Experiment 1, with the following exception. Participants were informed that they would be presented with an essay in which the words had been scrambled. We asked them to 'try to figure out what the essay was about' in order to answer questions about its content at the end of the experiment. Our intent was to keep the task as constant as possible across experiments (read for comprehension) while manipulating the presence or absence of syntactic structure in the stimulus set. As in Experiment 1, participants were asked to push a button after the prompt appeared when they were ready for the next series of words.

6. Results

6.1. Averages over word classes

Fig. 4A plots ERPs elicited by open- and closed-class words. Fig. 4B overplots ERPs from the left and right hemispheres for each word class. As in Experiment 1, a clear N1-P2 complex was visible in the first 300 ms. The N1 component peaked at about 120 ms. The P2 peaked at about 200 ms and, for both word class conditions, was larger over anterior sites ($F(4,52) = 9.25$, $P < 0.01$, MSe = 4.39) and larger over right hemisphere than left hemisphere sites, but only over posterior regions (hemisphere × electrode site: $F(4,52) = 4.25$, $P < 0.05$, MSe = 0.32).

Differences between word classes were in general not as robust as those observed in Experiment 1. As in Experiment 1, open-class words elicited a larger-amplitude P2 component (150–250 ms: $F(1,13) = 27.95$, $P < 0.001$, MSe = 1.28), and these differences between word class extended into the 250–350 ms window ($F(1,13) = 11.34$, $P < 0.01$, MSe = 3.54). Again as in Experiment 1, these differences tended to be larger at more posterior sites (word class × electrode site: $F(4,52) = 3.95$, $P = 0.05$, MSe = 0.34). Although a main effect for word class was not observed

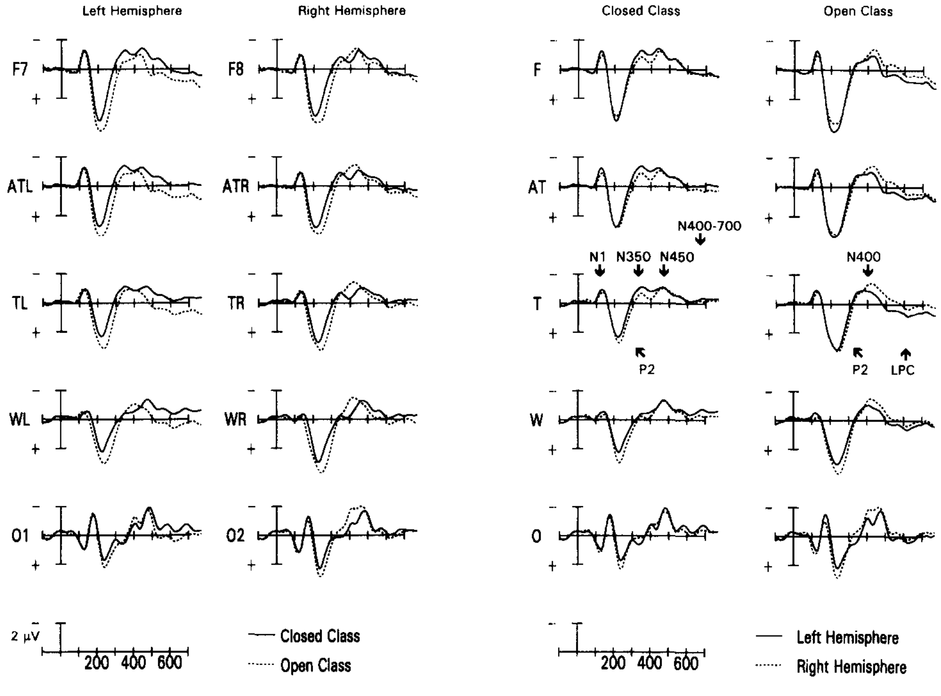

Fig. 4. Grand average ERPs recorded over 10 lateral sites to closed-class and open-class words in the scrambled prose condition. 1A overplots activity elicited by open- and closed-class words. 1B overplots activity recorded over the left and right hemispheres.

within the N400 window (350–450 ms), a reliable interaction between word class and hemisphere ($F(1,13) = 9.84$, $P < 0.01$, MSe = 0.69) reflected the fact that ERPs to closed-class words were more negative-going over the left hemisphere whereas the converse was true for open-class words. Within the 450 to 700 ms window, ERPs to closed-class words were more negative-going than were those to open-class words ($F(1,13) = 4.73$, $P < 0.05$, MSe = 4.09), but this difference was much larger in the left than in the right hemisphere (word class × hemisphere: $F(1,13) = 14.20$, $P < 0.03$, MSe = 0.42).

As in Experiment 1, closed-class words failed to elicit a distinct left-lateralized negativity at about 280 ms. Instead, the closed-class words elicited a slightly left-lateralized peak at about 350 ms and a second peak at about 450–500 ms. Open-class words elicited an N400 component that was slightly larger over the right hemisphere. Peak latencies and amplitudes for these negativities were computed as in Experiment 1. N400 latency was reliably later-occurring than N350 amplitude, ($F(1,13) = 56.38$, $P < 0.0001$, MSe = 0.003, and reliably earlier-occurring than N450 latency, $F(1,13) = 39.14$, $P < 0.0001$, MSe = 0.003. Differences in scalp topography of the N350 and N400 effects were also observed. The N350 was largest over the left hemisphere, whereas the N400 was largest over the right hemisphere (word class × hemisphere: $F(1,13) = 8.56$, $P < 0.02$, MSe = 0.87). No reliable distributional differences were found between the N450 and N400.

6.2. Averages over grammatical categories

ERPs averaged as a function of grammatical category are plotted in Fig. 5. Prepositions, pronouns, and auxiliaries elicited a negativity peaking at about 340 ms. Nouns and verbs elicited a negativity peaking at about 400 ms. Articles elicited a broad negative shift from about 280 ms to about 600 ms that was noticeably larger over the left hemisphere (N280–600). In each case, the amplitude of the negativity was largest over frontal and/or anterior temporal sites in the left hemisphere. The N400s elicited by nouns and verbs had maximum amplitudes over posterior, right hemisphere sites.

Statistical analyses were performed to determine whether these negativities differed in scalp distribution and latency, following the procedures specified for Experiment 1. Analyses on peak amplitude revealed reliable differences in scalp distribution across grammatical category (grammatical category × hemisphere: $F(5,65) = 5.16$, $P < 0.01$, MSe = 1.04; grammatical category × electrode site: $F(20,260) = 2.66$, $P = 0.06$, MSe = 1.12). Planned comparisons on peak amplitude were performed as in Experiment 1. No reliable differences were found in scalp distribution when comparing categories within the closed class. However, a reliable interaction between category and hemisphere was observed when comparing nouns and verbs, $F(1,13) = 6.21$, $P < 0.03$, MSe = 0.50. This interaction indicated that N400 amplitude was larger over the right than left hemisphere for both categories, whereas hemispheric differences were larger for nouns than for verbs. All three between-class comparisons revealed a reliable category by hemisphere interaction (articles vs. nouns: $F(1,13) = 15.46$, $P < 0.01$, MSe = 1.34; prepositions vs. nouns:

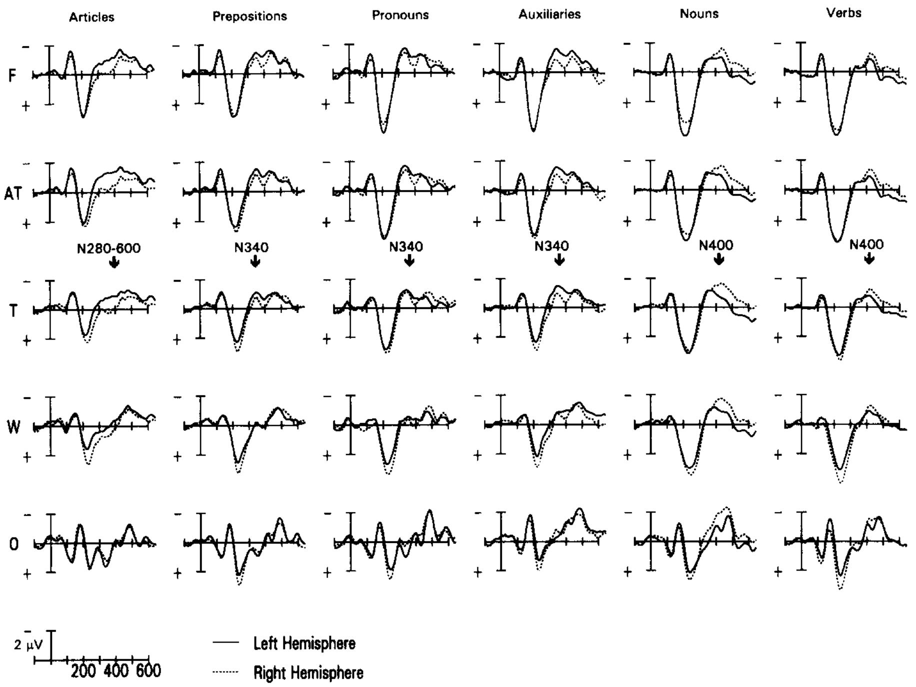

Fig. 5. ERPs averaged as a function of grammatical category for words in the scrambled prose condition.

$F(1,13) = 12.74$, $P < 0.01$, MSe = 1.02; pronouns vs. nouns: $F(1,13) = 13.76$, $P < 0.01$; MSe = 0.62). In all cases, these interactions indicated that the negativities to closed-class categories were larger over the left hemisphere at most sites, whereas the converse was true of the N400 elicited by nouns.

Differences in the latencies of the negativities across grammatical category were highly reliable, $F(5,65) = 67.87$, $P < 0.001$, MSe = 0.004. Regression analyses were conducted as in Experiment 1 in order to determine the relationship between latency and word length and frequency. Fig. 6 shows the resulting regression equations and scatterplots using the predictors of length and log frequency, respectively. As in Experiment 1, these correlations were highly reliable. Length and log frequency were again highly correlated with latency (length: $F(1,82) = 241.00$, $P < 0.0001$; frequency: $F(1,82) = 218.56$, $P < 0.0001$) and accounted for 75 and 73% of the variance, respectively. This was an increase of approximately 30% in the amount of variance accounted for in Experiment 1. The two-factor model accounted for 75% of the variance. The intercept, slope, and R^2-values for individual participants are shown in Table 2. In contrast to the results of Experiment 1, the between-subject (MSe = 0.0014) and within-subject (MSe = 0.0021) variance did not reliably differ.

As in Experiment 1, small differences in the anterior-posterior distribution of the N400–700 effect seemed to exist among grammatical categories from the closed class. An ANOVA on mean amplitude within the 450–700 ms window revealed a reliable interaction between category and electrode site (raw data: $F(12,156) = 3.00$, $P < 0.05$, MSe = 0.47; normalized data: $F(12,156) = 3.91$, $P < 0.01$).

6.3. *Normal vs. scrambled prose*

6.3.1. *Closed class*

In order to directly assess the effects of prose condition, we performed cross-experiment analyses comparing ERPs to the two word classes as a function of prose condition. Fig. 7A plots ERPs elicited in the normal and scrambled prose conditions by closed-class words. Closed-class words in the scrambled text condition elicited a slightly larger N1 component and were more negative between about 300 and 400 ms relative to the normal text condition, particularly over left frontal and midline frontal sites, and a slightly reduced N400–700, particularly between 500

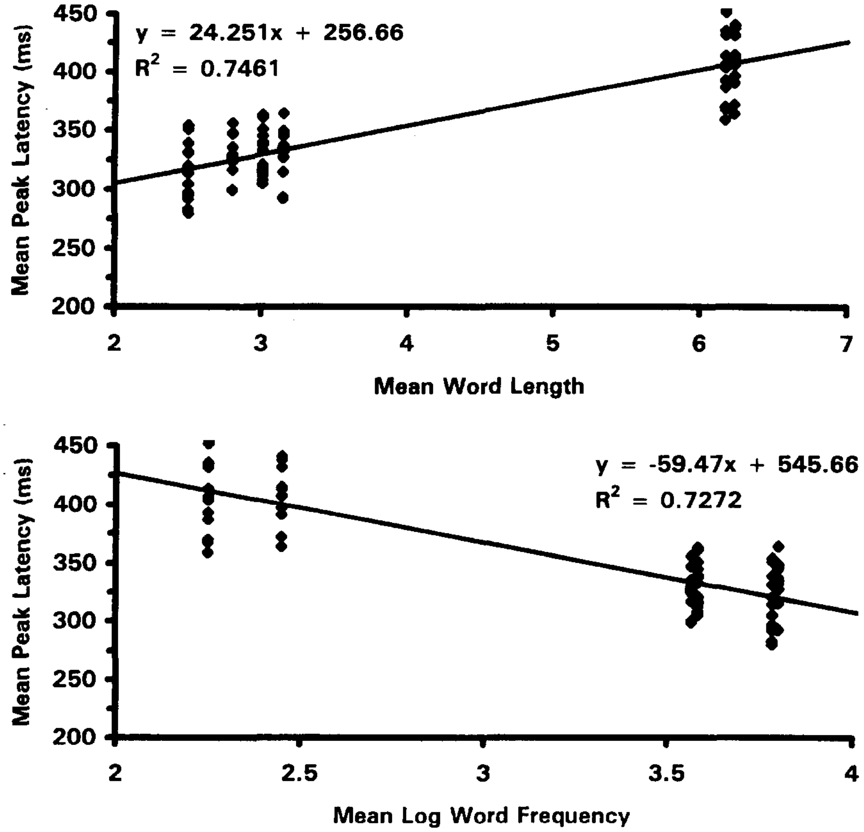

Fig. 6. Regression analyses correlating peak latency of the early-occurring negativity (in ms) with the mean length (top panel) and the log of mean normative frequency (bottom panel) of words in the scrambled prose condition.

Table 2
Individual regression parameters (ms) and proportion of variance accounted for, for mean peak latency regressed onto mean word length and mean log word frequency, scrambled text

Subject	Length			Log frequency		
	Intercept (ms)	Slope (ms)	R^2	Intercept (ms)	Slope (ms)	R^2
1	258	22	0.97	531	−56	0.97
2	231	24	0.94	519	−59	0.93
3	262	30	0.99	618	−74	0.99
4	291	18	0.69	492	−40	0.59
5	302	12	0.54	441	−29	0.53
6	232	31	0.87	600	−76	0.88
7	316	13	0.91	477	−33	0.91
8	249	22	0.84	505	−52	0.75
9	234	31	0.95	605	−76	0.93
10	240	34	0.91	636	−71	0.84
11	282	17	0.88	491	−43	0.92
12	197	35	0.93	621	−88	0.98
13	224	31	0.99	585	−74	0.92
14	275	20	0.98	519	−50	0.96

and 700 ms. To examine prose-condition effects between 300 and 400 ms, ANOVAs were performed on mean amplitude and peak amplitude treating prose condition as a between subjects factor. No reliable effects involving the group factor were observed. In particular, the two-way interaction between prose condition and electrode site and the three-way interaction between prose condition, hemisphere, and electrode site were unreliable ($P > 0.1$ in all analyses). An analysis on mean amplitude encompassing differences between 550 and 700 ms revealed that ERPs elicited in the scrambled condition were less negative-going than those elicited in the normal condition $F(1,26) = 4.53$, $P < 0.05$.

6.3.2. Open class

Fig. 7B plots ERPs elicited by open-class words. Relative to the normal prose condition, words in the scrambled condition elicited a slightly (and unreliably) enhanced N1 component. However, the most notable effect of prose condition was an increase in N400 amplitude in the scrambled condition relative to the normal condition, particularly over posterior sites (peak amplitude between 350 and 450 ms, prose condition × electrode site: $F(4,104) = 2.70$, $P < 0.05$, MSe $= 1.57$). Additionally, open-class words in the scrambled prose condition elicited a smaller late positive component between 450 and 700 ms than did the same words in the normal prose condition, $F(1,26) = 11.71$, $P < 0.01$, MSe $= 4.46$.

6.4. Discussion

As in Experiment 1, averages formed over word class showed that closed-class words failed to elicit an easily discerned N280 component. Instead, closed-class words elicited negative peaks at about 350 and 450 ms, whereas open-class words elicited an N400. However, in contrast to Experiment 1, the N350 and N400 had distinct hemispheric asymmetries. The N350 was larger over the left hemisphere whereas the N400 was larger over the right hemisphere.

Averages formed over grammatical categories again revealed that each category elicited a negative-going effect that immediately followed the N1-P2 complex and, further, that the latency of this negativity was predicted by word length and frequency. This result suggests that these early-occurring negativities might reflect some aspect of lexical processing rather than, e.g. syntactic analysis associated with sentence processing. A particularly striking finding was that the amount of latency variance accounted for by word length and frequency was approximately 30% greater in the scrambled prose condition than in the normal prose condition. This difference probably reflected the presence of greater between-subject variance in the normal prose condition. Again as in Experiment 1, differences in scalp distribution were found between grammatical categories, both for categories within one word class and for categories in different word classes.

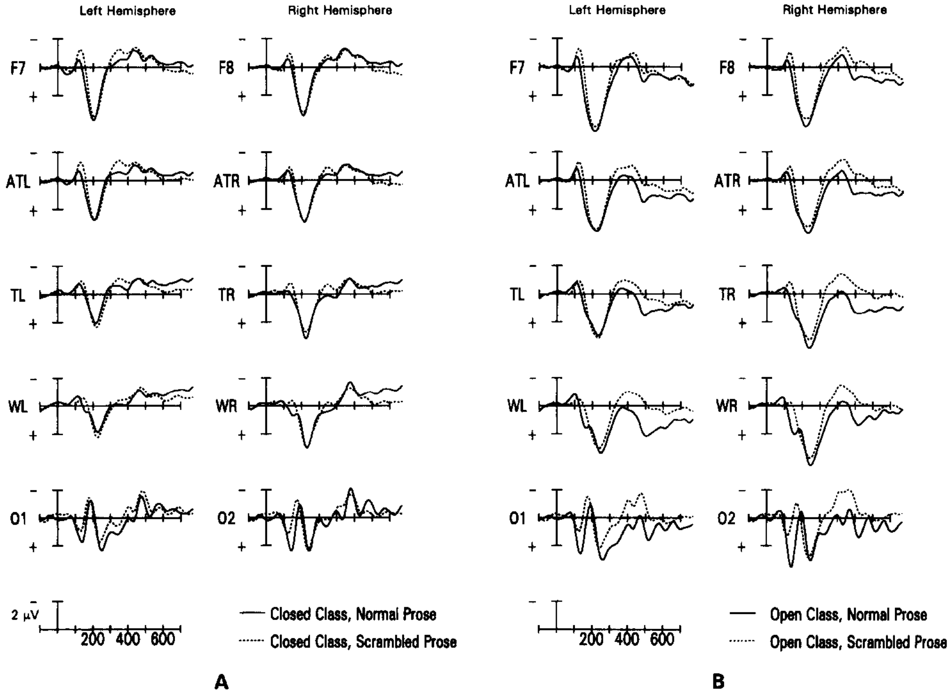

Fig. 7. ERPs elicited in the two prose conditions. (A) Closed-class words. (B) Open-class words.

Direct comparisons of ERPs to words in the normal and scrambled prose conditions revealed that open-class words elicited a larger-amplitude N400 in the scrambled prose than in the normal prose condition. This finding is consistent with prior work (e.g. Van Petten and Kutas, 1991) demonstrating that open-class words in coherent text elicit smaller N400s than do words in scrambled and consequently incoherent text. Furthermore, the N400–700 elicited by closed-class words and the LPC elicited by open-class words were reduced in amplitude in the scrambled prose condition, relative to the normal prose condition.

7. General discussion

In two experiments, we examined ERPs to words elicited during the processing of normal prose (Experiment 1) or scrambled prose (Experiment 2). The primary goal was to evaluate the claim that the N280 component is a categorical marker of the closed class and, more generally, that negative deflections in the ERP waveform reflect qualitative differences in the responses to the two word classes. The results reported here are largely inconsistent with these claims. ERPs averaged as a function of grammatical category showed that each category elicited an initial negativity following the N1-P2 complex. Critically, the peak latencies of these negativities were highly correlated with quantitative differences in word length and frequency. This was true both when participants read normal prose (Experiment 1) and when they read scrambled prose (Experiment 2). Prose condition had three notable effects. Firstly, the amount of latency variance accounted for by word frequency and length was greater in the scrambled condition than in the normal prose condition, probably due to the presence of greater between-subject variance in the normal prose condition. Secondly, open-class words elicited larger N400s when they appeared in scrambled prose than when they appeared in normal prose. This finding is consistent with a large body of evidence relating N400 amplitude to the ease with which a word can be semantically integrated with preceding context (e.g. Kutas and Hillyard, 1980a). Thirdly, the amplitudes of late-occurring effects (N400–700 for closed-class words and LPC for open-class words) were reduced in the scrambled prose condition, relative to the normal prose condition.

Our conclusion that the latency of these negativities is a function of word frequency and length, rather than word class, is consistent with the results of prior studies that have equated frequency and length across word class (Garnsey, 1985; Pulvermüller et al., 1995). These studies have found no latency differences in the ERPs to open- and closed-class words, a result that would be predicted assuming that frequency and length underlie this latency shift. Less consistent with this conclusion, however, are the results of prior studies examining frequency effects on the response to open-class words. These studies have shown that the frequency of open-class items affects N400 *amplitude*, but not N400 *latency* (Besson et al., 1992; Rugg, 1990). Clearly, more work is needed to resolve this issue. Conceivably, the frequency effects on latency are evident only within the higher frequency domains or with a sufficient range in frequencies. One goal of future research might be to carefully examine frequency effects by sampling a wide spectrum of frequencies. If this is done independently for both word classes, the slope of the regression line correlating these factors with the latency of the negativity could be directly compared for the two word classes. Diverging slopes would suggest that these negativities are in fact distinct across word class. Unfortunately, we sampled too few open-class grammatical categories to examine this possibility with the present data.

We should also reiterate that in the present study frequency and length categories were derived by collapsing across grammatical category and computing mean length and frequency for words in each category. As a result, the factors of frequency and length are perfectly confounded with grammatical category. Furthermore, the naturally occurring length and frequency variance within each category (and the fact that the frequency and length ranges of words within each category undoubtedly overlapped across categories) probably introduced noise into the correlations reported here. It would be useful to form categories directly as a function of length and frequency such that the boundaries between categories are nonoverlapping and have better-controlled variances. Under such conditions, one would anticipate even larger correlations than those observed here, assuming that the latency variation is indeed a function of word length and frequency.

Interestingly, the frequency effects on ERP latency reported here match up remarkably well with frequency effects on gaze duration during reading. For example, Just and Carpenter (1980) reported a clear relation between gaze duration and the logarithm of normative word frequency. Gaze duration increased by 53 ms for each log unit of decrease in word frequency. This compares with increases of 43 (Experiment 1) and 59 (Experiment 2) ms in the latency of the negativity reported here. It is also worth noting that a latency shift in a similar ERP negativity has been observed in the response to spoken words. Woodward et al. (1990) found that the latency of an N400-like negativity was highly correlated with the 'recognition point' of the spoken word (Marslen-Wilson and Tyler, 1980).

Three ERP effects observed here did seem to be a function of word class. As in earlier reports (Kutas and Hillyard, 1983; Kutas et al., 1988; Neville et al., 1992; St. George et al., 1994), closed-class words elicited an increased negativity between 400 and 700 ms, whereas open-class words elicited a late positive component (LPC) in this window. These effects were smaller in the scrambled-prose condition than in the normal-prose condition. Similar results were reported by Van Petten and Kutas (1991), who observed that N400–700 amplitude was largest for closed-class words in well-formed, nonanomalous sentences, and reduced in amplitude when the same words were embedded in random strings of words. Also consistent with this finding is the fact that the N400–700 seems to be absent when words are presented in a list format for lexical decision (Garnsey, 1985). These observations suggest that the N400–700 might reflect differences in how the two word classes are used during sentence or discourse comprehension rather than differences in how the word classes are represented or retrieved during word processing (c.f. Kutas and Van Petten, 1994). Finally, articles were the only grammatical category to elicit an effect that was quite similar to the N280 reported by Neville et al. (1992), but they did so only in the normal prose condition. Additional research is needed to more carefully examine the effects of grammatical category on ERPs.

It is perhaps appropriate to offer two cautionary notes that apply to evaluations of word-class differences in ERPs, whether reported here or elsewhere. Firstly, the effects of word class might in many cases be confounded with other variables that are correlated with word class. For example, prior work has revealed distributional effects attributable to imageability (e.g. Kounios and Holcomb, 1994), a factor that is highly correlated with word class. The obvious caveat is that one should be cautious in attributing causal status to the word class factor until other potentially confounding possibilities have been ruled out. A second cautionary note, discussed above, relates to evaluations of distributional and hemispheric asymmetries associated with the early-occurring negativities. Such evaluations are made somewhat problematic due to the potential effects of component overlap. In both experiments reported here, the P2 component was larger in amplitude over the right than the left hemisphere for both word classes and all grammatical categories. In most cases this asymmetry did not resolve prior to onset of the subsequent negative component. Consequently, it is difficult to discern whether hemispheric asymmetries within the time window of interest are due to asymmetries associated with the P2 component or to asymmetries associated with the subsequent negative-going component. Note that this problem of interpretation increases in severity with increasing temporal proximity between the P2 and the subsequent negativity. Indeed, the distributional effects observed in the present study might be largely (or entirely) due to overlap between the right-lateralized P2 and the subsequent negativity, particularly since the distance between these components (and thus the contaminating effects of overlap) differed across grammatical categories.

Comparisons of ERPs elicited by the various grammatical categories revealed a basic similarity, with a few notable exceptions in addition to the 'migrating' negativity. Most notably, the initial negative peak following the N1-P2 complex exhibited different left-right distributions across grammatical categories. This difference might be wholly or partly due to the effects of component overlap, i.e. the temporal proximity of this negativity with the immediately preceding, right-lateralized P2 component. Furthermore, the N400–700 effect showed distributional differences along the anterior-posterior dimension.

On a final note, King and Kutas (1995) have very recently reported findings that are remarkably similar to those reported here. King and Kutas found that ERPs to open- and closed-class words (presented in sentences) elicited similarly distributed negativities, and that the peak latencies of these negativities were a function of word length and frequency. Hence, the primary findings reported here and by King and Kutas seem to be robust.

In summary, quantitative differences in word frequency and length correlated highly with latency changes in an early-occurring negativity. These correlations were found both when words appeared in normal prose and when they appeared in scrambled prose, but were most robust in the scrambled prose condition. Furthermore, the negativities had similar scalp distributions across word type, and the distributional effects that were observed were a function of grammatical category rather than word class. These findings are seemingly inconsistent with the claim that open- and closed-class words elicit negativities that are qualitatively, reliably, and robustly distinct in distribution and timing.

Acknowledgements

Richard McKinnon is now at the University of Massachusetts, Amherst, MA. We thank two anonymous reviewers and Mireille Besson for very helpful comments on a previous version of this paper and Judy McLaughlin for help in running the experiment. This research was supported by research grant number 5 R29 DC 01947 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health awarded to the first author, and from a University of Washington Graduate Student Fund grant. A preliminary report of these data was presented at the Annual Meeting of the Psychonomic Society, 1994, St. Louis.

References

Besson, M., Kutas, M., and Van Petten, C. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience*, 4, 132–149.

Bradley, D.C. and Garrett, M.F. (1983). Hemispheric differences in the recognition of closed- and open-class words. *Neuropsychologia*, 21, 155–159.

Brown, C.M., and Hagoort, P. (1993). The processing nature of the N400: evidence from masked priming. *Journal of Cognitive Neuroscience*, 5, 34–44.

Caramazza, A. and Zurif, E.B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: evidence from aphasia. *Brain and Language*, 3, 572–582.

Coles, M.G.H., and Rugg, M.D. (1995). Event-related Brain Potentials: An Introduction. In: M.D. Rugg and M.G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford: Oxford University Press.

Damasio, H., and Damasio, A.R. (1989). *Lesion Analysis in Neuropsychology*. New York: Oxford University Press.

Flores d'Arcais, G.B. (1985). Children's recognition of function words. In G.A.J. Hoppebrouwers, P.A.M. Seuren, and A.J.M.M. Weitjers (Eds.), *Meaning and the Lexicon* (pp. 124–131). Dordrecht, the Netherlands: Foris Publications.

Francis, W N., and Kucera, H. (1982). *Computational Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin.

Friederici, A.D. (1985). Levels of processing and vocabulary types: evidence from online comprehension in normals and agrammatics. *Cognition*, 19, 133–166.

Garnsey, S.M. (1985). *Function words and content words: Reaction time and evoked potential measures of word recognition*. (Cognitive Science Tech. Rep. No. URCS-29). Rochester, NY: University of Rochester.

Garrett, M.F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language Production Vol. I Speech and Talk*. New York: Academic.

Garrett, M.F. (1982). A perspective on research in language production. In J. Mehler, E.C.T. Walker, and M. Garrett (Eds.), *Perspectives on mental representation*. Hillsdale, NJ: Erlbaum.

Greenhouse, S., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112.

Heilman, K.M., and Scholes, R.J. (1976). The nature of comprehension errors in Broca's conduction, and Wernicke's aphasics. *Cortex*, 12, 258–265.

Holcomb, P.J. (1993). Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. *Psychophysiology*, 30, 47–61.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Jarema, G., and Friederici, A.D. (1994). Processing articles and pronouns in agrammatic aphasia: Evidence from French. *Brain and Language*, 46, 683–694.

Jasper, H.H. (1958). Report to the committee on methods of clinical examination in electroencaphalography: the ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10, 371–375.

Johnson, R., Jr. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology*, 10, 90–97.

Just, M.A., and Carpenter, P.A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, 87, 329–354.

Kenan, L.R. (1986). *New American profiles*. New York: Harcourt Brace Jovanovitch.

Keppel, G. (1982). *Design and analysis: a researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.

King, J.W., and Kutas, M. (1995). The lexical processing negativity: an ERP whose latency indexes lexical characteristics of words. *Psychophysiology*, 35, S45.

Kounios, J., and Holcomb, P.J. (1994). Concreteness effects in semantic priming: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning Memory, and Cognition*, 20, 804–823.

Kutas, M. and Hillyard, S.A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, 11, 99–116.

Kutas, M. and Hillyard, S.A. (1980b). Reading between the lines: event-related brain potentials during natural sentence processing. *Brain and Language*, 11, 354–373.

Kutas, M. and Hillyard, S.A. (1980c). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, 207, 203–205.

Kutas, M. and Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, 11, 539–550.

Kutas, M., and Van Petten, C.K. (1994). Psycholinguistics electrified: event-related brain potential investigations. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego: Academic Press.

Kutas, M., Van Petten, C., and Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencehalography and Clinical Neurophysiology*, Sup. 39, 325–330.

Marslen-Wilson, W., and Tyler, L. (1980). The temporal structure of spoken-language understanding. *Cognition*, 8, 1–71.

McCarthy, G., and Wood, C.C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, 12, 203–208.

Neville, H.J., Mills, D.J., and Lawson, D.S. (1992). Fractionating language: different neural subsystems with different sensitive periods. *Cerebral Cortex* 2, 244–258.

Nobre, A.C., and McCarthy, G. (1994). Language-related ERPs: scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience*, 6, 233–255.

Osterhout, L., and Holcomb, P.J. (1995). Event-related potentials and language comprehension. In M.D. Rugg and M.G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford: Oxford University Press.

Pulvermüller, F., Lutzenberger, W., and Birbaumer, N. (1995). Electrocortical distinction of vocabulary types. *Electroencephalography and Clinical Neurophysiology*, 94, 357–370.

Rugg, M.D. (1990). Event-related brain potentials dissociate repetition effects of highand low-frequency words. *Memory and Cognition*, 17, 367–379.

Saffran, E.M., Schwartz, M.F., and Marin, O.S.M. (1980). The word order problem in agrammatism. II: Production. *Brain and Language*, 10, 249–262.

Schwartz, M., Saffran, E., and Marin, O. (1980). The word-order problem in agrammatism, I. Comprehension. *Brain and Language*, 10, 263–280.

St. George, M., Mannes, S., and Hoffinan, J.E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive Neuroscience*, 6, 70–83.

Van Petten, C. and Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory and Cognition*, 19, 95–112.

Woodward, S.H., Owens, J., and Thompson, L. W (1990). Word-to-word variation in ERP component latencies: Spoken words. *Brain and Language*, 38, 488–503.

Zurif, E., and Swinney, D.A. (1994). The neuropsychology of language. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego: Academic Press.